Online Resource 1
Supplementary Material for
“Variance-Calibrated Adjusted Two-Sample Blockwise Empirical Likelihood for the Difference of Means”

Reinis Alksnis Janis Valeinis

Laboratory of Statistical Research and Data Analysis,

Department of Mathematics, Faculty of Science and Technology,

University of Latvia, Jelgava Street 3, LV-1004 Riga, Latvia

Corresponding author, Reinis Alksnis, `reinis.alksnis@lu.lv`

For *Statistical Papers*

Abstract

This Supplement gives the assumptions and proofs for the higher-order, sample-based correction, and fixed-block-count results in the main article. Symbolic cross-checks, complete numerical results, and application sensitivity calculations are retained in the accompanying public GitHub repository.

1 Notation and proof structure

The notation follows the main paper. For sample $a = 1, 2$, $N_a = Q_a M_a$, $N = N_1 + N_2$, $Q = Q_1 + Q_2$, $\theta_a = Q_a/Q$, $\tau_a = N_a/N$, $\mathbf{M} = (M_1, M_2)$, $M_- = \min(M_1, M_2)$, and $M_+ = \max(M_1, M_2)$. In general $\theta_a \neq \tau_a$ when $M_1 \neq M_2$.

Section 4 records the profiled two-sample Bartlett specialization. Section 5 separates block and full-sample variance scales. Sections 6–8 extend the calculation to weak dependence and identify the adjacent-block term. Sections 9 and 10 treat estimated corrections. Section 11 controls the adjusted profile and proves its higher-order equivalence with multiplicative Bartlett correction. Section 12 gives the fixed-block-count theory. The repository contains symbolic cross-checks and longer computer-algebra output.

2 Global assumptions

Conditions S1–S4 below unpack the structural parts of (A1), (A2), and (A4) in the main paper. Conditions P1–P2 are the additional local profile conditions incorporated in (A3), while E1–E6 give the sparse-array formulation underlying its Edgeworth expansion.

2.1 Structural assumptions

S1. Sampling structure.

The two series are independent and strictly stationary with means μ_a. The sample sizes satisfy $N_a = Q_a M_a$, $M_a \to \infty$, $Q_a \to \infty$, and $\tau_{a,N} = N_a/N$ is bounded away from zero and one. For some fixed $0 < \kappa < 1/2$ and $0 < c_- < c_+ < \infty$,

$$c_- N^\kappa \leq M_a \leq c_+ N^\kappa, \qquad a = 1, 2.$$

Hence M_+/M_- is bounded, $Q_a \asymp Q \asymp N^{1-\kappa}$, and for some $\eta > 0$, $\eta \leq \theta_{a,N} = Q_a/Q \leq 1 - \eta$. We abbreviate the finite-N fractions to τ_a and θ_a.

S2. Covariance structure.

For each sample,

$$\sum_{h \geq 1} h^2 |\gamma_a(h)| < \infty, \qquad 0 < \underline{\sigma}^2 \leq \sigma_a^2 \leq \overline{\sigma}^2 < \infty.$$

S3. Weighted cumulant summability.

For $3 \leq r \leq 12$,

$$\sum_{h_2, \ldots, h_r \in \mathbb{Z}} \{1 + \operatorname{span}(0, h_2, \ldots, h_r)\} \left|\operatorname{cum}(Y_{a,0}, Y_{a,h_2}, \ldots, Y_{a,h_r})\right| < \infty.$$

S4. High moments.

For some $\delta > 0$ and $M_0 \geq 1$,

$$E|Y_{a,0}|^{30+\delta} < \infty, \qquad \sup_{m \geq M_0} E\left|\frac{\sum_{t=1}^m Y_{a,t}}{\{m\Omega_{a,m}\}^{1/2}}\right|^{30+\delta} < \infty.$$

2.2 Profile assumptions

P1. Local profile existence.

With probability $1 - o(N^{-1})$, the origin lies in the interior of the relevant sample convex hulls uniformly over the local nuisance-mean neighborhood used in profiling.

P2. Local profile regularity.

On the logarithmic neighborhood used below, the empirical second-moment scale is bounded away from zero and the local score denominators stay uniformly away from zero for $|t| + |d| \leq CQ^{-1/2}(\log Q)^K$. The empirical moments needed through degree six are uniformly bounded by a power of $\log Q$.

2.3 Probability assumptions

The sparse-array conditions E1–E6 are stated in Section 6. Conditional Cramér is retained separately from the mixing assumptions.

B1. Augmented Bartlett-moment expansion.

Jointly with the signed-root coordinates, the empirical second–fourth central-moment vector satisfies an order-six Borel–Edgeworth expansion and

$$\widehat{\eta}_N - \eta_N = Q^{-1/2} S_N + Q^{-1}\{\beta_{\eta,N} + \widetilde{D}_N\} + o_p(Q^{-1})$$

uniformly on the same logarithmic neighborhood, with the moment and cumulant bounds used in Section 9.

B2. Bartlett boundary smoothing.

The joint signed measure in B1 has a uniformly bounded smooth density in one root coordinate, and the boundary perturbation generated by $\widehat{b}_N^{\mathrm{blk}} - b_N^{\mathrm{blk}}$ admits a uniform second-order expansion with remainder $o(Q^{-2})$.

C1. Lag-window influence expansion.

For the full sample and each consecutive half sample, the centered lag-window estimators of σ_a^2 and B_{2a} admit joint first-order influence representations, uniformly for $h \leq L$, jointly with the signed-root coordinates.

C2. Second-order calibration bias.

The expectations of the full- and half-sample calibration estimators have a common-target second-order expansion. The complete homogeneous first-order inverse-sample-size bias doubles when both series are halved. The remaining deterministic terms are bounded by

$$O\Big\{\mathcal{T}_{\mathbf{M}}(L) + L^3/(M_- N^2) + r_{\mathbf{M}}\Big\}.$$

The centered H_2 projection used later is derived from S3 and C1 in Section 10.

C3. Joint boundary smoothing.

The joint full/half-sample influence vector in C1 satisfies the analogues of E2–E5 needed for the order-six expansion, and the resulting compactness/positivity event for $\widehat{\nu}_N^{JK}$ has complement probability $o(N^{-1})$.

3 Finite-block covariance identities

For one series, let M be its block length and $S_j = \sum_{t=(j-1)M+1}^{jM} Y_t$. Pair counting gives

$$\begin{aligned}
\Omega_M &= M^{-1}\operatorname{var}(S_1) && (1)\\
&= \gamma(0) + 2\sum_{h=1}^{M-1}\left(1-\frac{h}{M}\right)\gamma(h) && (2)\\
&= \sigma^2 - \frac{B_2}{M} + 2\sum_{h\geq M}\left(\frac{h}{M}-1\right)\gamma(h). && (3)
\end{aligned}$$

The final term is $o(M^{-2})$ under $\sum_h h^2|\gamma(h)| < \infty$. For adjacent blocks,

$$\operatorname{cov}(S_1, S_2) = \sum_{h=1}^{M} h\gamma(h) + \sum_{h=M+1}^{2M-1}(2M-h)\gamma(h) = \frac{B_2}{2} + o(M^{-1}). \tag{4}$$

Thus, with $Z_j = S_j/(M\Omega_M)^{1/2}$,

$$\operatorname{cov}(Z_1, Z_2) = \frac{B_2}{2M\sigma^2} + O(M^{-2}) = \frac{\delta}{2M} + O(M^{-2}).$$

For blocks separated by at least one complete block,

$$\sum_{k\geq 2}|\operatorname{cov}(S_0, S_k)| = o(M^{-1}), \qquad \sum_{k\geq 2}|\operatorname{cov}(U_0, U_k)| = o(M^{-3}),$$

and hence $\sum_{k\geq 2} |\operatorname{cov}(Z_0, Z_k)| = o(M^{-2})$.

4 Using the Liu et al. two-sample Bartlett correction

For fixed (M_1, M_2) the block means form an ordinary scalar two-sample empirical-likelihood problem. We therefore use the profiled two-sample Bartlett result of Liu et al. (2008) with the exact finite-block moments. The calculations below record the first signed-root grades because the same coordinates are needed when weak block dependence is added. They are not presented as a new derivation of the independent two-sample Bartlett theorem.

4.1 Profile recursion and first signed-root terms

Under H_0, shift sample 2 by Δ_0, translate the common nuisance mean to zero, put $q = Q^{-1/2}$, and define $\widehat{m}_{a,r}(d) = Q_a^{-1} \sum_{j=1}^{Q_a} (U_{a,j} - d)^r$. The profiled multiplier condition gives $\lambda_1 = t/\theta_1$ and $\lambda_2 = -t/\theta_2$, so

$$0 = \sum_{k\geq 0} (-t/\theta_1)^k \widehat{m}_{1,k+1}(d), \qquad 0 = \sum_{k\geq 0} (t/\theta_2)^k \widehat{m}_{2,k+1}(d).$$

Write

$$\widehat{m}_{a,r}(0) = m_{a,r} + qX_{a,r}, \qquad t = \sum_{j\geq 1} q^j t_j, \qquad d = \sum_{j\geq 1} q^j d_j,$$

and $\pi_a = (m_{a,2}/\theta_a)/H_{\mathbf{M}}$, $\pi_1 + \pi_2 = 1$. At grade j, after lower grades are inserted, $-d_j - \pi_1 t_j + A_j = 0$ and $-d_j + \pi_2 t_j + B_j = 0$, hence

$$t_j = A_j - B_j, \qquad d_j = \pi_2 A_j + \pi_1 B_j. \tag{5}$$

Lemma 4.1 (Local profile expansion)**.** *Under S1–S4 and P1–P2, the local profiled solution is unique on the logarithmic neighborhood for all sufficiently large N. For every fixed grade needed here, t and d are obtained by the recursion in* (5)*. The corrected signed root has the resulting polynomial expansion through grade Q^{-2} with remainder $O\{Q^{-5/2}(\log Q)^K\}$.*

Proof. At $q = 0$ and $(t, d) = (0, 0)$, the Jacobian of the two score equations with respect to (t, d) is

$$\begin{pmatrix} -m_{1,2}/\theta_1 & -1 \\ m_{2,2}/\theta_2 & -1 \end{pmatrix},$$

whose determinant is $H_{\mathbf{M}}$ and is uniformly separated from zero. Condition P2 keeps the empirical Jacobian nonsingular and the score denominators inside their analytic region. The analytic implicit-function theorem then gives a unique local solution. Taylor expansion in q gives the coefficient recursion above. The moment bounds in P2 control the derivatives on the logarithmic set and give the stated remainder after the fourth signed-root grade. $\square$

Normalize temporarily to $H_{\mathbf{M}} = 1$ and put

$$A = X_{1,1} - X_{2,1}, \qquad D_0 = \pi_2 X_{1,1} + \pi_1 X_{2,1},$$

$$V = \frac{X_{1,2}}{\theta_1} + \frac{X_{2,2}}{\theta_2}, \quad G = \frac{m_{1,3}}{\theta_1^2} - \frac{m_{2,3}}{\theta_2^2}, \quad T = \frac{X_{1,3}}{\theta_1^2} - \frac{X_{2,3}}{\theta_2^2}, \quad K = \frac{m_{1,4}}{\theta_1^3} + \frac{m_{2,4}}{\theta_2^3}.$$

Then $t_1 = A$ and $t_2 = -AV + GA^2$. To make the second grade transparent, put

$$\widetilde{X} = \frac{X_{1,1}}{\theta_1} + \frac{X_{2,1}}{\theta_2}, \qquad c_\pi = \frac{\pi_1}{\theta_1} - \frac{\pi_2}{\theta_2}.$$

The shifted second- and third-moment combinations satisfy

$$\widehat{H}(d) = 1 + qV + q^2 H_2 + O_p(q^3), \qquad H_2 = -2D_0\widetilde{X} + \frac{D_0^2}{\theta_1\theta_2},$$

$$\widehat{G}(d) = G + qG_1 + O_p(q^2), \qquad G_1 = T - 3D_0 c_\pi.$$

Direct expansion of the log profile first gives

$$w_2 = A^2V^2 - 2GA^3V + G^2A^4 - A^2H_2 + \frac{2}{3}A^3G_1 - \frac{1}{2}KA^4.$$

The identity

$$\widetilde{X} - Ac_\pi = \frac{D_0}{\theta_1\theta_2}$$

then reduces the two D_0 contributions exactly and yields

$$w_2 = A^2V^2 + \frac{A^2D_0^2}{\theta_1\theta_2} + \frac{2}{3}A^3T - 2GA^3V + \left(G^2 - \frac{K}{2}\right)A^4.$$

Thus the compact expression below is valid without requiring $\pi_1 = \theta_1$ or equal block lengths. Writing

$$W_N = A^2 + qw_1 + q^2w_2 + O_p(q^3), \qquad R_N = A + qP_1 + q^2P_2 + O_p(q^3),$$

coefficient matching gives

$$P_1 = \frac{G}{3}A^2 - \frac{1}{2}AV, \tag{6}$$

$$P_2 = A\left\{\frac{3}{8}V^2 + \frac{D_0^2}{2\theta_1\theta_2}\right\} + A^2\left\{\frac{1}{3}T - \frac{5}{6}GV\right\} + A^3\left\{\frac{4}{9}G^2 - \frac{1}{4}K\right\}. \tag{7}$$

4.2 Bartlett coefficient

Under the independent-block reference law,

$$E(X_{a,r}X_{a,s}) = \frac{m_{a,r+s} - m_{a,r}m_{a,s}}{\theta_a},$$

with cross-sample moments factorizing. Define $S = \pi_1^2/\theta_1 + \pi_2^2/\theta_2$ and $C = \pi_1\pi_2/(\theta_1\theta_2)$. The required contractions are

$$E(A^3) = qG, \quad E(A^2V) = q(K - S), \quad \mathrm{Var}(A) = 1,$$

$$\mathrm{Cov}(A, V) = G, \quad \mathrm{Var}(V) = K - S, \quad \mathrm{Var}(D_0) = \pi_1\pi_2, \quad \mathrm{Cov}(A, T) = K.$$

Substitution of (6)–(7) into $E\{H_2(R_N)\}$ gives

$$E\{H_2(R_N)\} = \frac{1}{Q}\left(-\frac{1}{3}G^2 + \frac{1}{2}K + C\right) + O(Q^{-2}). \tag{8}$$

For H_4, the Q^{-1} coefficient is

$$(K - 3S) + \{\tfrac{4}{3}G^2 - 6K + 6S\} + \{-\tfrac{4}{3}G^2 + 5K - 3S\} = 0,$$

so

$$E\{H_4(R_N)\} = O(Q^{-2}). \tag{9}$$

The three groups are the fourth cumulant of A, the first non-Gaussian correction involving P_1, and the leading Gaussian contractions involving P_2 and P_1^2. The remaining even Hermite components at order Q^{-1} cancel by the Bartlett-correctability result of Liu et al. (2008), applied here to the exact fixed-block distribution. The displayed contractions identify the factor and show explicitly where the nuisance-profile interaction enters.

Restoring scale gives

$$b_N^{\mathrm{blk}} = -\frac{G_{\mathbf{M}}^2}{3H_{\mathbf{M}}^3} + \frac{K_{\mathbf{M}}}{2H_{\mathbf{M}}^2} + \frac{C_{\mathbf{M}}}{H_{\mathbf{M}}^2}. \tag{10}$$

Combining the identified coefficient with the Bartlett-correctability theorem of Liu et al. (2008) gives

$$P(R_N^2 \le x) = G_1(x) - \frac{b_N^{\mathrm{blk}}}{Q}xg_1(x) + O(Q^{-2}), \tag{11}$$

and

$$P\left\{\frac{R_N^2}{1+b_N^{\mathrm{blk}}/Q} \le x\right\} = G_1(x) + O(Q^{-2}). \tag{12}$$

The independent-block correction is therefore imported from the two-sample result rather than rederived from first principles. Section 6 supplies the uniform triangular-array approximation needed below.

4.3 Role of the exact block moments

Let $\lambda_{3,a,M_a} = m_{a,3}/m_{a,2}^{3/2}$ and $\lambda_{4,a,M_a} = m_{a,4}/m_{a,2}^2 - 3$. Under cumulant summability, they are $O(M_a^{-1/2})$ and $O(M_a^{-1})$. Since $m_{a,2} = \theta_a \pi_{a,\mathbf{M}} H_{\mathbf{M}}$,

$$\frac{G_{\mathbf{M}}}{H_{\mathbf{M}}^{3/2}} = \frac{\lambda_{3,1,M_1}\pi_{1,\mathbf{M}}^{3/2}}{\sqrt{\theta_1}} - \frac{\lambda_{3,2,M_2}\pi_{2,\mathbf{M}}^{3/2}}{\sqrt{\theta_2}},$$

$$\frac{K_{\mathbf{M}}}{H_{\mathbf{M}}^2} = \sum_{a=1}^{2} \frac{\{3+\lambda_{4,a,M_a}\}\pi_{a,\mathbf{M}}^2}{\theta_a}, \qquad \frac{C_{\mathbf{M}}}{H_{\mathbf{M}}^2} = \frac{\pi_{1,\mathbf{M}}\pi_{2,\mathbf{M}}}{\theta_1\theta_2}.$$

Hence $b_N^{\mathrm{blk}} = b_N^G + O(M_-^{-1})$ and its non-Gaussian part contributes $O(Q^{-1}M_-^{-1}) = O(N^{-1})$. It cannot be discarded before applying the exact block-moment correction.

5 Exact population variance calibration

5.1 Block and full-sample variance scales

Let $Y_{a,t} = X_{a,t} - \mu_a$ and $U_{a,j} = M_a^{-1}\sum_{t=(j-1)M_a+1}^{jM_a} Y_{a,t}$. Since $N_a = Q_a M_a$, $\bar{U}_a = \bar{Y}_a$. Put

$$D_N = \bar{Y}_1 - \bar{Y}_2. \tag{13}$$

For a stationary sample of length n, define

$$\Omega_{a,n} = \frac{1}{n}\,\mathrm{Var}\left(\sum_{t=1}^{n} Y_{a,t}\right). \tag{14}$$

Then $v_{a,M_a} = \mathrm{Var}(U_{a,j}) = \Omega_{a,M_a}/M_a$, and

$$V_N = \frac{\Omega_{1,N_1}}{N_1} + \frac{\Omega_{2,N_2}}{N_2}, \tag{15}$$

$$V_{\mathbf{M}} = \frac{\Omega_{1,M_1}}{N_1} + \frac{\Omega_{2,M_2}}{N_2}. \tag{16}$$

5.2 Exact quadratic-profile identity

The leading quadratic profile is

$$\mathcal{Q}_N(m) = \frac{Q_1(\bar{U}_1 - m)^2}{v_{1,M_1}} + \frac{Q_2(\bar{U}_2 - m)^2}{v_{2,M_2}}. \tag{17}$$

With $a_a = Q_a/v_{a,M_a}$, its minimizer is $\widehat{m}_0 = (a_1\bar{U}_1 + a_2\bar{U}_2)/(a_1 + a_2)$, and

$$\begin{aligned} W_{N,0} := \inf_m \mathcal{Q}_N(m) &= \frac{a_1 a_2}{a_1 + a_2}(\bar{U}_1 - \bar{U}_2)^2 \\ &= \frac{D_N^2}{v_{1,M_1}/Q_1 + v_{2,M_2}/Q_2}. \end{aligned} \tag{18}$$

Since $v_{a,M_a}/Q_a = \Omega_{a,M_a}/N_a$,

$$W_{N,0} = \frac{D_N^2}{V_{\mathbf{M}}}. \tag{19}$$

Define

$$\nu_N = \frac{V_{\mathbf{M}}}{V_N}. \tag{20}$$

Then

$$\nu_N W_{N,0} = \frac{D_N^2}{V_N}. \tag{21}$$

This is an exact finite-sample identity.

In the one-sample scalar specialization, the same identity gives

$$\frac{\Omega_N}{\Omega_M} = 1 + \frac{B_2}{M\sigma^2} + o(M^{-1}) + O(N^{-1}).$$

Thus, when $B_2 \neq 0$, the leading uncalibrated statistic contains an M^{-1} scale error. A Bartlett-type modification whose leading effect is of order Q^{-1} cannot remove this term. At $M \asymp N^{1/3}$, the scale error is $N^{-1/3}$. This is the calculation behind the rate comparison with Kitamura (1997) and the scalar specialization of Wang and Polonik (2019) discussed in the main paper. It concerns higher-order accuracy and not the first-order chi-square limit.

5.3 Scope of the cancellation

Equation (21) removes the complete block-to-full-sample variance-scale term from the leading quadratic profile. For the full statistic,

$$\nu_N W_N = \frac{D_N^2}{V_N} + \nu_N(W_N - W_{N,0}), \tag{22}$$

so profile nonlinearities, block dependence, higher block cumulants, and their interactions remain.

5.4 Exact covariance identity and approximate calibration

With $\sigma_a^2 = \sum_{h\in\mathbb{Z}} \gamma_a(h)$ and $B_{2a} = 2\sum_{h=1}^{\infty} h\gamma_a(h)$, pair counting gives

$$\Omega_{a,n} = \gamma_a(0) + 2\sum_{h=1}^{n-1}\left(1 - \frac{h}{n}\right)\gamma_a(h) = \sigma_a^2 - \frac{B_{2a}}{n} + r_{a,n}, \tag{23}$$

where

$$r_{a,n} = 2\sum_{h\geq n}\left(\frac{h}{n} - 1\right)\gamma_a(h). \tag{24}$$

Consequently,

$$V_{\mathbf{M}} = \sum_{a=1}^{2} \frac{1}{N_a}\left(\sigma_a^2 - \frac{B_{2a}}{M_a} + r_{a,M_a}\right), \tag{25}$$

$$V_N = \sum_{a=1}^{2} \frac{1}{N_a}\left(\sigma_a^2 - \frac{B_{2a}}{N_a} + r_{a,N_a}\right). \tag{26}$$

Under $\sum_h h^2|\gamma_a(h)| < \infty$, $r_{a,M_a} = o(M_a^{-2})$, which need not be $o(N^{-1})$ at $M_a \asymp N^{1/3}$. Thus $r_{\mathbf{M}} = o(N^{-1})$ in the sample-based theorem is an additional covariance-tail condition and is not implied by S2. The population expansion uses the exact ratio (20) and does not require this stronger condition.

5.5 Higher cumulants after exact standardization

Under full-sample cumulant summability, for $r \geq 3$,

$$\mathrm{cum}_r\left(\sum_{t=1}^{N_a} Y_{a,t}\right) = O(N_a), \qquad \mathrm{cum}_r(\bar{Y}_a) = O(N_a^{1-r}).$$

Independence and balanced sample proportions give $\mathrm{cum}_r(D_N) = O(N^{1-r})$ and $V_N \asymp N^{-1}$. Thus, for $Z_N = D_N/\sqrt{V_N}$,

$$\mathrm{cum}_r(Z_N) = O\{N^{-(r-2)/2}\}. \tag{27}$$

The third- and fourth-order terms are $N^{-1/2}$ and N^{-1}. The odd $N^{-1/2}$ contribution cancels for the symmetric event $Z_N^2 \leq x$.

6 Uniform Edgeworth approximation for the block arrays

This section supplies the probability approximation used to turn the signed-root algebra into uniform CDF expansions. We invoke Lahiri (2007, Theorem 2.1). E1–E6 are our adapted conditions.

For sample a, let

$$Z_{a,j,M_a} = \frac{\sum_{t=(j-1)M_a+1}^{jM_a} Y_{a,t}}{\sqrt{M_a \Omega_{a,M_a}}}, \qquad L_{a,N} = Q_a^{-1/2} \sum_{j=1}^{Q_a} Z_{a,j,M_a},$$

and for $s \geq 5$ define

$$P_{a,N}^{(s)} = Q_a^{-1/2} \sum_{j=1}^{Q_a} \left(Z_{a,j,M_a}^2 - EZ_{a,1,M_a}^2, \ldots, Z_{a,j,M_a}^{s-1} - EZ_{a,1,M_a}^{s-1} \right)',$$

with $T_{a,N}^{(s)} = (L_{a,N}, P_{a,N}^{(s)\prime})'$. In Lahiri's notation, use $n = N_a$, $\ell = M_a$, $b = Q_a$, $X_{i,n} = Y_{a,i}$, and the sparse block coordinate

$$\mathcal{Y}_{i,n} = \begin{cases} M_a h_{s-1,M_a}^{+}(Z_{a,j,M_a}), & i = (j-1)M_a + 1, \\ 0, & \text{otherwise.} \end{cases}$$

Then $T_{a,N}^{(s)}$ is a deterministic uniformly nonsingular transform of the joint normalized sum.

We impose the following conditions uniformly for both samples.

(E1) $M_a, Q_a \to \infty$, $Q_a/Q \in [\eta, 1-\eta]$, and $c_- N^\kappa \leq M_a \leq c_+ N^\kappa$ for some $0 < \kappa < 1/2$.

(E2) The moment and local-approximation bounds required for an order-s Borel expansion hold for the sparse joint array. For $s = 6$, S4 is a convenient sufficient moment condition.

(E3) Lahiri's local approximation and strong-mixing conditions hold uniformly, and the covariance matrices of the joint vectors remain uniformly nonsingular.

(E4) Lahiri's conditional Cramér condition holds for the joint linear–degree-$(s-1)$ block-power array.

(E5) The normalized joint cumulants entering the Edgeworth polynomials have the required uniform expansions through order $s - 2$.

(E6) On $\mathcal{G}_N = \{\|T^{(s)}_{1,N}\| \vee \|T^{(s)}_{2,N}\| \le (D\log Q)^{1/2}\}$, one linear coordinate of the polynomial signed-root map supplied by Lemma 4.1 has derivative bounded away from zero.

Proposition 6.1. *Under E1–E6, for the order-s polynomial approximation there is a Gaussian-polynomial signed measure* $\Psi^{(s)}_N$ *such that, uniformly in* $x \ge 0$*,*

$$\begin{aligned}\left|P\{(R^c_N)^2 \le x\} - \mathfrak{E}_{N,s}(x)\right| \le\, & CQ^{-(s-2)/2}(\log N)^{-2} \\ & + CQ^{-(s-1)/2}(\log Q)^K,\end{aligned} \tag{E.1}$$

where $\mathfrak{E}_{N,s}(x) = \Psi^{(s)}_N\{(u,v) : \mathcal{R}_{N,s}(u,v)^2 \le x\}$*. For* $s = 6$ *the right-hand side is* $o(N^{-1})$ *whenever* $0 < \kappa < 1/2$*.*

Proof. Apply Lahiri (2007, Theorem 2.1) separately to the two sparse arrays. The signed measures have Gaussian-polynomial densities with uniformly bounded weighted total variation. A smooth majorant of $1\{\|T^{(s)}_{a,N}\| > (D\log Q)^{1/2}\}$ gives $P(\mathcal{G}^c_N) = o(N^{-1})$ for D sufficiently large. Because the two samples are independent, the product of the two marginal signed measures gives the joint signed measure. The derivative condition in E6 gives one-coordinate anti-concentration for the squared-root boundary, including at $x = 0$. Lemma 4.1 gives the local root remainder, which changes the event only in a tube of width $O\{Q^{-(s-1)/2}(\log Q)^K\}$. Since $Q \asymp N^{1-\kappa}$, both error terms are $o(N^{-1})$ for $s = 6$ and $0 < \kappa < 1/2$. □

For $s = 6$, the preceding algebra gives

$$\mathfrak{E}_{N,6}(x) = G_1(x) + \mathcal{E}_N(x) + O(Q^{-2}) + o(N^{-1}).$$

Thus Proposition 6.1 yields the population theorem in the main article. At $\kappa = 1/2$, $Q^{-2} \asymp N^{-1}$ and all terms of order Q^{-2} must also be retained.

7 From Gaussian block calculations to general blocks

This section identifies the dependence channels that can contribute at order N^{-1} after the independent-block Bartlett correction. We use the exact finite-block weights $\pi_{a,\mathbf{M}}$.

7.1 Standardized block variables and terms that remain

Write

$$S_{a,j} = \sum_{t=(j-1)M_a+1}^{jM_a} Y_{a,t}, \qquad Z_{a,j} = \frac{S_{a,j}}{\{M_a\Omega_{a,M_a}\}^{1/2}}.$$

Then $EZ_{a,j} = 0$, $\operatorname{var}(Z_{a,j}) = 1$, and under S3,

$$\operatorname{cum}_r(Z_{a,j}) = O(M_a^{1-r/2}), \qquad 3 \le r \le 12. \tag{T.1}$$

In particular,

$$\lambda_{3,a,M_a} = O(M_a^{-1/2}), \qquad \lambda_{4,a,M_a} = O(M_a^{-1}). \tag{T.2}$$

The adjacent standardized covariance is

$$\rho_{a,M_a} = \operatorname{cov}(Z_{a,1}, Z_{a,2}) = \frac{B_{2a}}{2M_a\sigma_a^2} + O(M_a^{-2}) = \frac{\delta_a}{2M_a} + O(M_a^{-2}). \tag{T.3}$$

Its linear contribution is $Q^{-1}\rho_{a,M_a} = O(Q^{-1}M_a^{-1}) = O(N^{-1})$.

7.2 Non-Gaussian reduction of the adjacent-link coefficient

Assume S1–S4, P1–P2, and E1–E6. After the independent-block terms of order Q^{-1} have been removed using the exact block moments, every remaining term absent from the Gaussian one-adjacent-link calculation that contains an adjacent covariance with a marginal third or fourth cumulant, a cross-block cumulant of order at least three, or a covariance at block lag at least two contributes $o(Q^{-1}M_-^{-1})$ to the CDF of the squared corrected root. Hence the N^{-1} dependence coefficient is the linear adjacent-covariance coefficient obtained from the Gaussian block calculation.

Proof. From (T.2)–(T.3),

$$\rho_{a,M_a}\lambda_{3,a,M_a} = O(M_a^{-3/2}), \qquad \rho_{a,M_a}\lambda_{4,a,M_a} = O(M_a^{-2}), \qquad \rho_{a,M_a}\lambda_{3,a,M_a}^2 = O(M_a^{-2}). \tag{T.4}$$

After multiplication by Q^{-1} these are $o(Q^{-1}M_-^{-1})$.

Weighted cumulant summability gives

$$\operatorname{cum}(Z_{a,1}, Z_{a,1}, Z_{a,2}) = O(M_a^{-3/2}), \tag{T.5}$$

$$\operatorname{cum}(Z_{a,1}, Z_{a,1}, Z_{a,2}, Z_{a,2}) = O(M_a^{-2}). \tag{T.6}$$

A third-order cross-block cumulant appears at an odd signed-root grade and cancels under the even projection unless paired with another odd term. Its first surviving products are therefore $O(M_a^{-2})$ or smaller. The fourth-order term is already $O(M_a^{-2})$.

For more distant blocks,

$$\sum_{k\geq 2} |\operatorname{cov}(Z_{a,1}, Z_{a,1+k})| = o(M_a^{-2}). \tag{T.7}$$

The relevant orders can be summarized as

channel	size before the CDF projection	status
$Q^{-1}\rho_a$	$Q^{-1}M_a^{-1}$	retained
$Q^{-1}\rho_a\lambda_{3,a}$	$Q^{-1}M_a^{-3/2}$	$o(N^{-1})$
$Q^{-1}\rho_a\lambda_{4,a}$	$Q^{-1}M_a^{-2}$	$o(N^{-1})$
$Q^{-1}\rho_a\lambda_{3,a}^2$	$Q^{-1}M_a^{-2}$	$o(N^{-1})$
$Q^{-1}\rho_a^2$	$Q^{-1}M_a^{-2}$	$o(N^{-1})$
$Q^{-1}\operatorname{cum}_4(1,1,2,2)$	$Q^{-1}M_a^{-2}$	$o(N^{-1})$
$Q^{-1}\sum_{k\geq 2}\|\operatorname{cov}(Z_1, Z_{1+k})\|$	$o(Q^{-1}M_a^{-2})$	$o(N^{-1})$.

A third-order cross-block cumulant is odd at its first signed-root grade and therefore has no linear even projection. Its first surviving product is of order $Q^{-1}M_a^{-2}$ or smaller. Pure marginal third- and fourth-cumulant terms belong to the exact finite-block independent-reference Bartlett correction and are removed there. Higher cross-block cumulants have at least the same additional boundary-tail penalty under E2–E5. Thus the only N^{-1} dependence term is linear in (T.3), giving

$$-xg_1(x)\left[\frac{\pi_{1,\mathbf{M}}\delta_1}{2\theta_2 N_1}\{2\theta_2 - L_{\mathbf{M}}x\} + \frac{\pi_{2,\mathbf{M}}\delta_2}{2\theta_1 N_2}\{2\theta_1 - L_{\mathbf{M}}x\}\right] = \mathcal{E}_N(x).$$

□

8 Gaussian one-link identification of the adjacent-block coefficient

By scale invariance set $H_{\mathbf{M}} = 1$, write $\theta_1 = \theta$, $\theta_2 = 1-\theta$, $\pi_1 = \pi$, $\pi_2 = 1-\pi$, and define

$$J_{\mathbf{M}} = \frac{\pi^2}{\theta} + \frac{(1-\pi)^2}{1-\theta}, \qquad \widetilde{C}_{\mathbf{M}} = \frac{\pi(1-\pi)}{\theta(1-\theta)}, \qquad L_{\mathbf{M}} = \theta(1-\theta)J_{\mathbf{M}}.$$

Let ρ_a be the lag-one covariance of the standardized Gaussian block sequence in sample a, all other inter-block covariances being zero, and put

$$J_\rho = \frac{\pi^2\rho_1}{\theta} + \frac{(1-\pi)^2\rho_2}{1-\theta}, \quad \Pi_\rho = \pi\rho_1 + (1-\pi)\rho_2, \quad \Xi_\rho = (1-\pi)\rho_1 + \pi\rho_2.$$

The population VC factor satisfies

$$\nu_N^{-1} = 1 + 2\sum_{a=1}^{2} \pi_{a,\mathbf{M}} \sum_{k=1}^{Q_a-1} \left(1 - \frac{k}{Q_a}\right) \rho_a(k). \tag{G.1}$$

For a covariance perturbation Σ_ε and $F_\varepsilon = H_j(R_N^c)$,

$$\frac{d}{d\varepsilon} E_\varepsilon F_\varepsilon \bigg|_0 = \frac{1}{2}\sum_{r,s} \dot{\Sigma}_{rs} E_0(\partial_r\partial_s F_0) + E_0\big(\partial_\varepsilon F_\varepsilon\big|_0\big)\,. \tag{G.2}$$

The second term is required because R_N^c contains ν_N. The $Q_a^{-1}\rho_a$ row-end term in (G.1) cancels in the leading variance and is only $O(Q^{-2}\rho)$ when inserted into a term already carrying Q^{-1}.

Let $A_a = X_{a,1}$, $V_a = X_{a,2}/\theta_a$ and, to first order in ρ,

$$v_a = \pi_a(1+2\rho_a), \qquad \sigma_A^2 = 1 + 2\Pi_\rho, \qquad \tau_a^2 = 2\pi_a^2/\theta_a.$$

Projecting V_a on A_a gives $V_a = q(A_a^2 - v_a)/\theta_a + \widetilde{V}_a$, with $\widetilde{V}_a$ independent of A_a at the retained order and $\mathrm{var}(\widetilde{V}_a) = \tau_a^2 + O(q^2+\rho^2)$. Hence

$$\begin{aligned}
E(A^2V) &= 2q\sum_a \frac{v_a^2}{\theta_a}, & E(A^4V) &= 12q\sigma_A^2\sum_a \frac{v_a^2}{\theta_a},\\
E(A^2V^2) &= \sigma_A^2\sum_a \tau_a^2, & E(A^4V^2) &= 3\sigma_A^4\sum_a \tau_a^2,\\
E(A^3T) &= 9\sigma_A^2\sum_a \frac{m_{a,2}v_a}{\theta_a^2}, & E(A^5T) &= 45\sigma_A^4\sum_a \frac{m_{a,2}v_a}{\theta_a^2}.
\end{aligned} \tag{G.3}$$

Also,

$$\begin{aligned}
&\mathrm{var}(D_0) = \pi(1-\pi)(1+2\Xi_\rho) + O(Q^{-1}\rho + \rho^2),\\
&\sum_a \frac{v_a^2}{\theta_a} = J_{\mathbf{M}} + 4J_\rho + O(\rho^2), \qquad \sum_a \frac{m_{a,2}v_a}{\theta_a^2} = J_{\mathbf{M}} + 2J_\rho + O(\rho^2),\\
&\sum_a \tau_a^2 = 2J_{\mathbf{M}}.
\end{aligned} \tag{G.4}$$

For Gaussian marginal blocks $G = 0$, $K = 3J_{\mathbf{M}}$, and $b = 3J_{\mathbf{M}}/2 + \widetilde{C}_{\mathbf{M}}$. The reduction preceding (7) shows that the second-order coefficient is exactly $w_2 = A^2V^2 + A^2D_0^2/\{\theta(1-\theta)\} + 2A^3T/3 - KA^4/2$. Substitution gives

$$E(W_N) = 1 + 2\Pi_\rho + \frac{b}{Q} + \frac{4J_\rho - 2J_{\mathbf{M}}\Pi_\rho + 2\widetilde{C}_{\mathbf{M}}(\Pi_\rho + \Xi_\rho)}{Q} + o(Q^{-1}\rho) + O(\rho^2), \tag{G.5}$$

$$E(W_N^2) = 3 + 12\Pi_\rho + \frac{6b}{Q} + \frac{24J_\rho - 6J_{\mathbf{M}}\Pi_\rho + 24\widetilde{C}_{\mathbf{M}}\Pi_\rho + 12\widetilde{C}_{\mathbf{M}}\Xi_\rho}{Q} + o(Q^{-1}\rho) + O(\rho^2), \tag{G.6}$$

$$E(W_N^3) = 15 + 90\Pi_\rho + \frac{45b}{Q} + \frac{180J_\rho + 270\widetilde{C}_{\mathbf{M}}\Pi_\rho + 90\widetilde{C}_{\mathbf{M}}\Xi_\rho}{Q} + o(Q^{-1}\rho) + O(\rho^2). \tag{G.6a}$$

Using

$$E\{H_2(R_N^c)\} = \frac{\nu_N E(W_N)}{1 + b/Q} - 1, \qquad E\{H_4(R_N^c)\} = \frac{\nu_N^2 E(W_N^2)}{(1+b/Q)^2} - 6\frac{\nu_N E(W_N)}{1+b/Q} + 3,$$

and the analogous identity for H_6, the $O(\rho)$ scale terms cancel. Equations (G.5)–(G.6a) also give

$$E\{H_6(R_N^c)\}_{\mathrm{lin}} = o(Q^{-1}\rho) + O(\rho^2) + O(Q^{-2}),$$

so no sixth-Hermite term survives in the retained adjacent-block calculation. The remaining terms are

$$E\{H_2(R_N^c)\}_{\mathrm{lin}} = \frac{1}{Q}\{4J_\rho + 2\widetilde{C}_{\mathbf{M}}\Xi_\rho - 5J_{\mathbf{M}}\Pi_\rho\}, \tag{G.7}$$

$$E\{H_4(R_N^c)\}_{\mathrm{lin}} = -\frac{12}{Q}J_{\mathbf{M}}\Pi_\rho. \tag{G.8}$$

Writing $c_a = \theta_a \pi_{a,\mathbf{M}}\rho_a/Q$ gives

$$E\{H_2(R_N^c)\}_{\mathrm{lin}} = \frac{c_1}{\theta^2(1-\theta)}\{2(1-\theta) - 3L_{\mathbf{M}}\} + \frac{c_2}{(1-\theta)^2\theta}\{2\theta - 3L_{\mathbf{M}}\}, \tag{G.9}$$

$$E\{H_4(R_N^c)\}_{\mathrm{lin}} = -12L_{\mathbf{M}}\left\{\frac{c_1}{\theta^2(1-\theta)} + \frac{c_2}{(1-\theta)^2\theta}\right\}. \tag{G.10}$$

Finally,

$$P\{(R_N^c)^2 \le x\} - G_1(x) = -xg_1(x)\left[E\{H_2(R_N^c)\} + \frac{x-3}{12}E\{H_4(R_N^c)\}\right].$$

Since $\rho_a = \delta_a/(2M_a) + O(M_a^{-2})$ and $N_a = \theta_a Q M_a$, (G.9)–(G.10) reduce to

$$-xg_1(x)\left[\frac{\pi_{1,\mathbf{M}}\delta_1}{2\theta_2 N_1}\{2\theta_2 - L_{\mathbf{M}}x\} + \frac{\pi_{2,\mathbf{M}}\delta_2}{2\theta_1 N_2}\{2\theta_1 - L_{\mathbf{M}}x\}\right] + o(N^{-1}),$$

which is $\mathcal{E}_N(x)$. The repository symbolic calculation is retained only as an independent cross-check of (G.7)–(G.10).

9 Estimation of the Bartlett factor

Put $q = Q^{-1/2}$ and let η_N collect the second–fourth central block moments entering $B(\eta_N) = b_N^{\mathrm{blk}}$. For sample a, define

$$X_a = \sqrt{Q}\,\bar{U}_a, \qquad S_{a,r} = \sqrt{Q}\,Q_a^{-1}\sum_{j=1}^{Q_a}\psi_{a,r}(U_{a,j}),$$

where

$$\psi_{a,2}(u) = u^2 - m_{a,2}, \quad \psi_{a,3}(u) = u^3 - m_{a,3} - 3m_{a,2}u, \quad \psi_{a,4}(u) = u^4 - m_{a,4} - 4m_{a,3}u.$$

The empirical central moments satisfy

$$\begin{aligned}
\widehat{m}_{a,2} - m_{a,2} &= qS_{a,2} - q^2X_a^2 + o_p(q^2),\\
\widehat{m}_{a,3} - m_{a,3} &= qS_{a,3} - 3q^2X_aS_{a,2} + o_p(q^2),\\
\widehat{m}_{a,4} - m_{a,4} &= qS_{a,4} + q^2\{-4X_aS_{a,3} - 6m_{a,2}X_a^2\} + o_p(q^2).
\end{aligned} \tag{28}$$

Hence

$$\widehat{\eta}_N - \eta_N = qS_N + q^2\{\beta_{\eta,N} + \widetilde{D}_N\} + o_p(q^2), \qquad E_G\widetilde{D}_N = 0,$$

and a second-order delta expansion gives

$$\widehat{b}_N^{\mathrm{blk}} - b_N^{\mathrm{blk}} = qB_{1,N} + q^2B_{2,N} + o_p(q^2), \tag{29}$$

with

$$B_{1,N} = g_B'S_N, \qquad B_{2,N} = g_B'\beta_{\eta,N} + g_B'\widetilde{D}_N + \tfrac{1}{2}S_N'H_BS_N.$$

Let A be the leading standard-normal root contrast and $v = \mathrm{Cov}_G(S_N, A)$. The even conditional projection of $B_{2,N}$ on A is

$$E_G(B_{2,N} \mid A) = \beta_{b,0} + \beta_{b,2}^{\text{full}} H_2(A) + \{\text{odd terms}\},$$

where

$$\beta_{b,0} = g_B' \beta_{\eta,N} + \tfrac{1}{2} \operatorname{tr}(H_B \Sigma_\eta),$$

$$\beta_{b,2}^{\text{full}} = \tfrac{1}{2} v' H_B v + \tfrac{1}{2} E_G\{H_2(A) g_B' \widetilde{D}_N\}.$$

The last expectation is a finite Gaussian-moment contraction and is determined by the moments already entering B.

Put $\bar{G}_{\mathbf{M}} = G_{\mathbf{M}}/H_{\mathbf{M}}^{3/2}$ and $\rho_b = \operatorname{Cov}_G(A, B_{1,N})$. By scale invariance take $H_{\mathbf{M}} = 1$ for the following calculation. Since $R_N^c = A + qP_1 + O_p(q^2)$ with $P_1 = GA^2/3 - AV/2$, Wick contraction gives

$$m_R = E_G(P_1) = -G/6, \qquad \operatorname{cum}_G(P_1, A, A) = -G/3.$$

Let $\eta_b = \operatorname{Cov}_G(V, B_{1,N})$. The single-block cumulant identity gives

$$q^{-1} \operatorname{cum}(A, A, B_{1,N}) = \eta_b + o(1),$$

and Wick's rule gives

$$\operatorname{cum}_G(P_1, A, B_{1,N}) = -\eta_b/2 + G\rho_b/6.$$

Consequently

$$q^{-1} \operatorname{cum}(R_N^c, R_N^c, B_{1,N}) = G\rho_b/3 + o(1).$$

Restoring scale yields

$$m_R = -\bar{G}_{\mathbf{M}}/6, \qquad \kappa_{300} = 0, \qquad \kappa_{210} = \bar{G}_{\mathbf{M}} \rho_b/3. \tag{30}$$

Hence $\rho_b m_R + \kappa_{210}/2 = 0$. Under B2, the boundary expansion leaves only the quadratic projection of $B_{2,N}$ and gives

$$P(T_{b,N} \le x) - P(T_N \le x) = Q^{-2} x g_1(x) \{\beta_{b,0} + \beta_{b,2}^{\text{full}}(x-1)\} + o(Q^{-2}). \tag{31}$$

Hence estimating the Bartlett factor changes the CDF only at order Q^{-2}.

10 Sample-based variance calibration and jackknife

Define

$$\mathcal{V}_{N,\mathbf{M}}(s_1, b_1, s_2, b_2) = \frac{\sum_{a=1}^{2} N_a^{-1}(s_a - b_a/M_a)}{\sum_{a=1}^{2} N_a^{-1}(s_a - b_a/N_a)},$$

$$\widehat{\nu}_N^F = \mathcal{V}_{N,\mathbf{M}}(\widehat{\sigma}_1^2, \widehat{B}_{21}, \widehat{\sigma}_2^2, \widehat{B}_{22}),$$

and

$$\mathcal{T}_a(L) = \sum_{h>L} h|\gamma_a(h)|, \qquad \mathcal{T}_{\mathbf{M}}(L) = \max_a \mathcal{T}_a(L)/M_a.$$

Split each sample into consecutive halves and insert the half-sample lag-window estimates into the same full-target functional, holding N_a, M_a, L fixed. If the resulting estimates are $\widehat{\nu}_N^{(1)}$ and $\widehat{\nu}_N^{(2)}$, define

$$\widehat{\nu}_N^{JK} = 2\widehat{\nu}_N^F - \frac{1}{2}\{\widehat{\nu}_N^{(1)} + \widehat{\nu}_N^{(2)}\}. \tag{32}$$

Under C2,

$$E\widehat{\nu}_N^{JK} - \nu_N = O\left\{\mathcal{T}_{\mathbf{M}}(L) + \frac{L^3}{M_- N^2} + r_{\mathbf{M}}\right\}. \tag{33}$$

A first-order perturbation of the calibration ratio is

$$\delta\nu_N = \frac{\sum_a N_a^{-1}\{(1-\nu_N)e_a - f_a/M_a + \nu_N f_a/N_a\}}{\sum_a N_a^{-1}(\sigma_a^2 - B_{2a}/N_a)} + \text{quadratic remainder},$$

and $1 - \nu_N = O(M_-^{-1})$, so the covariance-tail errors are block-scale damped.

For

$$\mathcal{K}_a(h) = \sum_{r,s\in\mathbb{Z}} \operatorname{cum}(Y_{a,0}, Y_{a,h}, Y_{a,r}, Y_{a,s}),$$

S3 gives $\sum_h (1+|h|)|\mathcal{K}_a(h)| < \infty$. For one sample, let $Z_n = \sum_{t=1}^n Y_t/(n\Omega_n)^{1/2}$. Direct cumulant expansion of the centered sample autocovariance gives, uniformly for $h \le L$,

$$E[(\widehat{\gamma}_n(h) - E\widehat{\gamma}_n(h))H_2(Z_n)] = \frac{\mathcal{K}(h)}{n\sigma^2} + O\left(\frac{1+h}{n^2}\right).$$

The leading Gaussian pairings cancel because the subtraction generated by sample centering has the same first-order projection as the uncentered autocovariance. Summation over the lag window gives an $O(n^{-1}) + O(L^3 n^{-2})$ projection for the two lag-window functionals. The derivatives of $\mathcal{V}_{N,\mathbf{M}}$ with respect to these functionals are $O(M_a^{-1})$. The average of the two half-sample influence functions equals the full-sample influence to first order, so only boundary and separate-centering remainders remain. Therefore

$$E[(\widehat{\nu}_N^{JK} - E\widehat{\nu}_N^{JK})H_2(A_N)] = O\biggl(\frac{1}{M_- N} + \frac{L^3}{M_- N^2}\biggr). \tag{34}$$

Combining the population adjacent-link term, the Bartlett plug-in terms of order Q^{-2}, (33), (34), quadratic calibration perturbations, and covariance approximation gives

$$\begin{aligned}\sup_{x\geq 0}\Bigl|P(\widehat{W}_N \leq x) - G_1(x) - \mathcal{E}_N(x)\Bigr| = O\Bigl(Q^{-2} + \frac{1}{M_- N} + \mathcal{T}_{\mathbf{M}}(L) \\ + \frac{L^3}{M_-^2 N} + \frac{L^3}{M_- N^2} + r_{\mathbf{M}}\Bigr).\end{aligned} \tag{35}$$

A sufficient population-equivalence regime is

$$M_+ = o(N^{1/2}), \qquad L \to \infty, \qquad L^3 = o(M_-^2), \tag{36}$$

together with

$$\mathcal{T}_{\mathbf{M}}(L) = o(N^{-1}), \qquad r_{\mathbf{M}} = o(N^{-1}). \tag{37}$$

Then $P(\widehat{W}_N \leq x) = G_1(x) + \mathcal{E}_N(x) + o(N^{-1})$ uniformly in $x \geq 0$. The jackknife changes the deterministic bandwidth restriction from $L = o(M_-^{1/2})$ to

$$L = o(M_-^{2/3}). \tag{38}$$

If $|\gamma_a(h)| \leq C\rho^h$ and $M_a \asymp N^\kappa$, $L = c\log N$ satisfies the direct tail condition whenever $c\log(1/\rho) > 1 - \kappa$.

11 Adjusted two-sample BEL and higher-order equivalence

The pseudo-observation construction is the two-sample AEL form with the sample-specific block-count weights used in the main paper. The half-Bartlett principle is due to Liu and Yu (2010). Dependent higher-order AEL is treated by Ma (2014), Wang and Polonik (2019). Here we need the profiled two-sample specialization after variance calibration.

11.1 Construction and profile existence

Under H_0, put

$$\widetilde{U}_{1,j} = U_{1,j}, \qquad \widetilde{U}_{2,j} = U_{2,j} + \Delta_0, \qquad g_{a,j}(m) = \widetilde{U}_{a,j} - m,$$

and append $g_{a,0}(m;\alpha) = -\alpha\theta_a \bar{g}_a(m)$, $\alpha > 0$. The pseudo-value has the opposite sign to the original sample mean whenever that mean is nonzero. If the mean is zero, the original values already contain both signs unless all are zero. Hence zero belongs to the augmented scalar hull for every m and the adjusted profile is finite.

For fixed block counts, the adjusted-EL expansion of Liu and Yu (2010), with the sample-specific weights used here, gives

$$b_N^{\mathrm{blk},A}(\alpha) = b_N^{\mathrm{blk}} - 2\alpha, \qquad \alpha^*_{\mathbf{M}} = b_N^{\mathrm{blk}}/2. \tag{39}$$

Pearson's standardized moment inequality gives $b_N^{\mathrm{blk}} \geq 1/2$ and $\alpha^*_{\mathbf{M}} \geq 1/4$. The same argument applies to the empirical central moments whenever $\widehat{H}_{\mathbf{M}} > 0$.

11.2 Local perturbation and control of the full profile

Put $q = Q^{-1/2}$. At the ordinary optimizer,

$$t = qA + q^2(-AV + GA^2) + O_p(q^3),$$

while the pseudo-log contribution is

$$J = -2\alpha q^2 A^2 - 2\alpha q^3 A(-AV + GA^2) + O_p(q^4).$$

The profiled nuisance curvature is order Q, whereas the adjustment score is smaller, so re-optimization changes the criterion only at order q^4, uniformly for α in a fixed compact subset of $(0, \infty)$.

Let $\rho_Q = Q^{-1/2}(\log Q)^c$, $c > 1/2$, and $\delta_Q = Q^{-1/4}$. On the baseline moment event, local curvature gives

$$\inf_{\rho_Q < |m| \leq \delta_Q} \{W_N^A(m;\alpha) - W_N^A(\widehat{m}_A;\alpha)\} \geq c_1(\log Q)^{2c}.$$

For $\delta_Q < |m| \leq 4\Lambda_N$, let u be the uniform weights on the augmented sample and let $p^*(m)$ be the maximizing empirical-likelihood weights. Pinsker's inequality and the moment constraint give

$$W^A_{a,N}(m;\alpha) = 2(Q_a+1)D\{u\|p^*(m)\} \geq \frac{(Q_a+1)\bar{g}_{a,u}(m)^2}{G_a(m)^2},$$

where $G_a(m) = \max_j |g_{a,j}(m;\alpha)|$. On the moment event, $\Lambda_N \leq Q^{3/(30+\delta)}$, so this region is separated by a lower bound of order $Q^{1/2-6/(30+\delta)}$. For $|m| > 4\Lambda_N$, the moment constraint forces the pseudo-observation to receive a fixed positive fraction of the total mass. An AM–GM bound then gives $W^A_N(m;\alpha) \geq c_B Q - C\log Q$, uniformly for $\alpha \in [1/4, B]$. These barriers dominate the local profile value. Hence there is an event $\mathcal{L}^A_N$ such that

$$P\{(\mathcal{L}^A_N)^c\} = o(N^{-1}) \tag{40}$$

and the global adjusted minimizer lies in $|m| \leq \rho_Q$.

11.3 Signed-root equivalence for the sample-based statistics

On $\mathcal{L}^A_N$, analytic factorization at $A = 0$ gives $W_N = A^2H_N$ and $W^A_N = A^2H^A_N$. The local calculation yields

$$W^A_N(\alpha) - \frac{W_N}{1+2\alpha/Q} = -\frac{2\alpha}{3}\frac{G_{\mathbf{M}}}{H^{3/2}_{\mathbf{M}}}Q^{-3/2}A^3 + O\{Q^{-2}A^2(1+|A|)^K\}, \tag{41}$$

and

$$R^A_N(\alpha) - R^{BC}_N(\alpha) = -\frac{\alpha}{3}\frac{G_{\mathbf{M}}}{H^{3/2}_{\mathbf{M}}}Q^{-3/2}A^2 + O\{Q^{-2}|A|(1+|A|)^K\}. \tag{42}$$

Since $G_{\mathbf{M}}/H^{3/2}_{\mathbf{M}} = O(M_-^{-1/2})$, the right-hand side is $o(N^{-1})$ on this event for $M_a \asymp N^\kappa$, $0 < \kappa < 1/2$. Root-scale anti-concentration then gives an $o(N^{-1})$ CDF difference.

The empirical half-Bartlett adjustment is bounded below by $1/4$ and lies in a fixed compact interval with complement probability $o(N^{-1})$. The factor $\widehat{\nu}^{JK}_N$ is similarly bounded away from zero and infinity under (36)–(37). Comparing the two sample-based statistics with their common $\widehat{\nu}^{JK}_N$ and $\widehat{b}^{\text{blk}}_N$ gives

$$\sup_{x\geq 0}\left|P(\widehat{T}^A_N \leq x) - G_1(x) - \mathcal{E}_N(x)\right| = o(N^{-1}).$$

Thus VC+ABEL and VC+BC have the same displayed N^{-1} adjacent-block coefficient.

12 Fixed block counts and the profiled Gaussian reference law

This section uses a separate asymptotic regime from S1.

12.1 Fixed-block-count assumptions

FQ1. Sampling regime.

Fix $q_1, q_2 \geq 3$. Let $Q_a = q_a$, $M_a \to \infty$, $M_1/M_2 \to c \in (0, \infty)$, and define $\tau^2_{a,M_a} = \Omega_{a,M_a}/M_a$. Assume $r_N = \tau_{1,M_1}/\tau_{2,M_2} \to r \in (0, \infty)$. Under S2, $r = (\sigma_1/\sigma_2)c^{-1/2}$.

FQ2. Finite-vector block central limit theorem.

With $\widetilde{U}_{1,j} = U_{1,j}$ and $\widetilde{U}_{2,j} = U_{2,j} + \Delta_0$, put

$$Z_{a,j,N} = \frac{\widetilde{U}_{a,j} - \mu_1}{\tau_{a,M_a}}.$$

Assume

$$(Z_{1,1,N}, \ldots, Z_{1,q_1,N}, Z_{2,1,N}, \ldots, Z_{2,q_2,N}) \Rightarrow (Z_1, Z_2),$$

where all coordinates are independent standard Gaussian variables.

FQ3. Scale estimation.

Assume $\widehat{\sigma}^2_a \to_p \sigma^2_a$ and $\widehat{\nu}^{JK}_N \to_p 1$. Then

$$\widehat{r}_N = \left\{ \frac{\widehat{\sigma}^2_1/M_1}{\widehat{\sigma}^2_2/M_2} \right\}^{1/2} \to_p r.$$

12.2 Finite-dimensional profile functionals

For $z = (z_1, \ldots, z_q)$ and $u \in \mathbb{R}$, define

$$\ell_q(z, u) = 2 \sum_{j=1}^{q} \log\{1 + \lambda(z_j - u)\},$$

where λ solves $\sum_{j=1}^{q}(z_j - u)/\{1 + \lambda(z_j - u)\} = 0$ with positive denominators. Set $\ell_q(z, u) = +\infty$ outside the interior of the sample range. For $d > 0$, append $h_0 = -d(\bar{z} - u)$ to $h_j = z_j - u$

and define

$$\ell_q^A(z,u;d) = 2\sum_{j=0}^{q}\log(1+\lambda^A h_j),$$

where λ^A solves the augmented moment equation. The adjusted functional is finite for every finite u.

Its common limit as $u \to \pm\infty$ is

$$L_q(d) = 2\left\{(q+1)\log\frac{d+1}{q+1} + q\log\frac{q}{d}\right\}. \tag{43}$$

Thus ℓ_q^A extends continuously to the extended real line $\overline{\mathbb{R}}$, which permits a profile infimum at infinity.

12.3 Limit of the empirical half-Bartlett adjustment

Let $s_{a,k} = q_a^{-1}\sum_j (Z_{a,j} - \bar{Z}_a)^k$ and $\vartheta_a = q_a/(q_1+q_2)$. Since the first limiting block scale is normalized to one and the second equals $1/r$,

$$\widehat{m}_{1,k} = \tau_{1,M_1}^k s_{1,k,N}, \qquad \widehat{m}_{2,k} = \tau_{1,M_1}^k r_N^{-k} s_{2,k,N}.$$

The common factor cancels from the Bartlett formula. Write

$$H_r = \frac{s_{1,2}}{\vartheta_1} + \frac{r^{-2}s_{2,2}}{\vartheta_2}, \qquad G_r = \frac{s_{1,3}}{\vartheta_1^2} - \frac{r^{-3}s_{2,3}}{\vartheta_2^2},$$

$$K_r = \frac{s_{1,4}}{\vartheta_1^3} + \frac{r^{-4}s_{2,4}}{\vartheta_2^3}, \qquad C_r = \frac{r^{-2}s_{1,2}s_{2,2}}{\vartheta_1^2\vartheta_2^2},$$

and

$$\mathcal{B}_r = -\frac{G_r^2}{3H_r^3} + \frac{K_r}{2H_r^2} + \frac{C_r}{H_r^2}. \tag{44}$$

Then $\widehat{b}_N^{\mathrm{blk}} \Rightarrow \mathcal{B}_r(Z_1, Z_2)$ and $\mathcal{B}_r \geq 1/2$ almost surely.

12.4 Proof of the fixed-block-count limit

Under H_0, set $t = (m-\mu_1)/\tau_{1,M_1}$. Then

$$\widetilde{U}_{1,j} - m = \tau_{1,M_1}(Z_{1,j,N} - t), \qquad \widetilde{U}_{2,j} - m = \tau_{2,M_2}(Z_{2,j,N} - r_N t).$$

Scale invariance gives the exact representation

$$W_N(\Delta_0) = \inf_{t \in \mathbb{R}} \{\ell_{q_1}(Z_{1,N}, t) + \ell_{q_2}(Z_{2,N}, r_N t)\}.$$

The ordinary functional is continuous at Gaussian vectors except for endpoint coincidences, which have probability zero. Hence, on $[0, +\infty]$,

$$W_N(\Delta_0) \Rightarrow \mathcal{W}_{q_1,q_2}(r).$$

Jointly, $\widehat{b}_N^{\text{blk}} \Rightarrow \mathcal{B}_r$ and $\widehat{\nu}_N^{JK} \to_p 1$, so

$$\widehat{W}_N \Rightarrow \frac{\mathcal{W}_{q_1,q_2}(r)}{1 + \mathcal{B}_r/(q_1 + q_2)}.$$

For adjusted BEL, $\widehat{\alpha}_{\mathbf{M}} = \widehat{b}_N^{\text{blk}}/2 \Rightarrow \mathcal{B}_r/2$. The continuous extension in (43) makes the adjusted objective continuous jointly in the block vectors, r, adjustment, and nuisance coordinate, so

$$W_N^A(\Delta_0; \widehat{\alpha}_{\mathbf{M}}) \Rightarrow \mathcal{W}_{q_1,q_2}^A(r).$$

FQ3 and Slutsky's theorem give the same limit for $\widehat{T}_N^A$.

12.5 Ordinary profile failure for equal scales

When $r = 1$, ordinary BEL is finite exactly when the two Gaussian sample ranges overlap in their interiors. Failure occurs when all observations from one sample lie below all observations from the other. Since all $\binom{q_1+q_2}{q_1}$ label orderings are equally likely and exactly two separate the ranges,

$$P\{\mathcal{W}_{q_1,q_2}(1) = +\infty\} = \frac{2}{\binom{q_1+q_2}{q_1}}.$$

12.6 Plug-in critical values, reciprocal-scale symmetry, and Wilks compatibility

Let $F_{q_1,q_2,r}^A$ be the distribution function of $\mathcal{W}_{q_1,q_2}^A(r)$ and $c_{1-\gamma}(q_1, q_2, r)$ its quantile. Weak continuity of the reference law in $r > 0$ and FQ3 imply, when the quantile is unique and continuous,

$$c_{1-\gamma}(q_1, q_2, \widehat{r}_N) \to_p c_{1-\gamma}(q_1, q_2, r).$$

For equal block counts, $\vartheta_1 = \vartheta_2 = 1/2$ and direct substitution gives

$$H_{1/r}(Z_2, Z_1) = r^2 H_r(Z_1, Z_2), \qquad G_{1/r}(Z_2, Z_1) = -r^3 G_r(Z_1, Z_2),$$

$$K_{1/r}(Z_2, Z_1) = r^4 K_r(Z_1, Z_2), \qquad C_{1/r}(Z_2, Z_1) = r^4 C_r(Z_1, Z_2),$$

so $\mathcal{B}_{1/r}(Z_2, Z_1) = \mathcal{B}_r(Z_1, Z_2)$. After the change $s = rt$,

$$\mathcal{W}^A_{q,q}(r; Z_1, Z_2) = \mathcal{W}^A_{q,q}(1/r; Z_2, Z_1),$$

and therefore

$$\mathcal{W}^A_{q,q}(r) \stackrel{d}{=} \mathcal{W}^A_{q,q}(1/r), \qquad c_{1-\gamma}(q, q, r) = c_{1-\gamma}(q, q, 1/r).$$

Whenever differentiable at $r = 1$,

$$\frac{\partial}{\partial \log r} c_{1-\gamma}(q, q, r)\bigg|_{r=1} = 0.$$

Finally let $q_1, q_2 \to \infty$, with ϑ_a bounded away from zero and one and r in a compact subset of $(0, \infty)$. Then $s_{a,2} \to_p 1$, $s_{a,3} \to_p 0$, $s_{a,4} \to_p 3$, and

$$\mathcal{B}_r \to_p b_\infty(\vartheta_1, \vartheta_2, r),$$

where

$$b_\infty = \frac{\frac{3}{2}\{\vartheta_1^{-3} + r^{-4}\vartheta_2^{-3}\} + r^{-2}(\vartheta_1\vartheta_2)^{-2}}{\{\vartheta_1^{-1} + r^{-2}\vartheta_2^{-1}\}^2}.$$

This limit is bounded on compact parameter sets, so $d_a = \vartheta_a \mathcal{B}_r/2 = O_p(1)$ and $d_a/q_a = o_p(1)$. The adjusted-EL expansion of Chen et al. (2008) then gives, uniformly over the local nuisance region,

$$\ell^A_{q_1}(Z_1, t; d_1) + \ell^A_{q_2}(Z_2, rt; d_2) = q_1(\bar{Z}_1 - t)^2 + q_2(\bar{Z}_2 - rt)^2 + o_p(1).$$

The minimum of the quadratic part is

$$\frac{(\bar{Z}_1 - \bar{Z}_2/r)^2}{q_1^{-1} + r^{-2}q_2^{-1}},$$

which is exactly χ^2_1. Hence $\mathcal{W}^A_{q_1,q_2}(r) \Rightarrow \chi^2_1$. In the balanced case with $r = 1$, $b_\infty = 5/2$.

13 Causal linear-process and AR(1) formulas

For $Y_{a,t} = \sum_{j\geq 0} a_{a,j}\varepsilon_{a,t-j}$, let the $\varepsilon_{a,t}$ be i.i.d. mean-zero innovations with variance $\kappa_{2,a}$. If $\sum_j j^2|a_{a,j}| < \infty$ and $\mathcal{A}_a = \sum_j a_{a,j} \neq 0$, put $c_a(h) = \sum_j a_{a,j}a_{a,j+h}$. Then

$$\gamma_a(h) = \kappa_{2,a}c_a(h), \qquad \sigma_a^2 = \kappa_{2,a}\mathcal{A}_a^2, \qquad B_{2a} = 2\kappa_{2,a}\sum_{h\geq 1} hc_a(h).$$

Under the corresponding innovation moments, the coefficient condition gives the covariance and cumulant summability used above. The Edgeworth and conditional Cramér conditions remain separate.

For AR(1), $Y_{a,t} = \phi_a Y_{a,t-1} + \varepsilon_{a,t}$,

$$\gamma_a(h) = \frac{\kappa_{2,a}}{1-\phi_a^2}\phi_a^{|h|}, \qquad \sigma_a^2 = \frac{\kappa_{2,a}}{(1-\phi_a)^2}, \qquad \frac{B_{2a}}{\sigma_a^2} = \frac{2\phi_a}{1-\phi_a^2},$$

and

$$\frac{\Omega_{a,M_a}}{\sigma_a^2} = 1 - \frac{2\phi_a(1-\phi_a^{M_a})}{M_a(1-\phi_a^2)}, \qquad \operatorname{cov}(S_{a,1}, S_{a,2}) = \gamma_a(0)\frac{\phi_a(1-\phi_a^{M_a})^2}{(1-\phi_a)^2}.$$

Thus the adjacent standardized-block correlation is $B_{2a}/(2M_a\sigma_a^2) + O(M_a^{-2})$.

References

Jiahua Chen, Asokan Mulayath Variyath, and Bovas Abraham. Adjusted empirical likelihood and its properties. *Journal of Computational and Graphical Statistics*, 17(2):426–443, 2008. doi: 10.1198/106186008X321068.

Yuichi Kitamura. Empirical likelihood methods with weakly dependent processes. *The Annals of Statistics*, 25(5):2084–2102, 1997. doi: 10.1214/aos/1069362388.

Soumendra N. Lahiri. Asymptotic expansions for sums of block-variables under weak dependence. *The Annals of Statistics*, 35(3):1324–1350, 2007. doi: 10.1214/009053607000000190.

Yukun Liu and Chi Wai Yu. Bartlett correctable two-sample adjusted empirical likelihood. *Journal of Multivariate Analysis*, 101(7):1701–1711, 2010. doi: 10.1016/j.jmva.2010.03.008.

Yukun Liu, Changliang Zou, and Runchu Zhang. Empirical likelihood for the two-sample mean problem. *Statistics & Probability Letters*, 78(5):548–556, 2008. doi: 10.1016/j.spl.2007.09.006.

Jun Ma. *Essays on Empirical Likelihood*. PhD thesis, University of British Columbia, Vancouver, Canada, 2014. URL `https://library-archives.canada.ca/eng/services/services-libraries/theses/Pages/item.aspx?idNumber=1033224685`.

Guangxing Wang and Wolfgang Polonik. High order adjusted block-wise empirical likelihood for weakly dependent data, 2019. arXiv:1912.07699v3, last revised 13 August 2021.